A Total Perspective on Simulation Uncertainty

Dale K. Pace, Consultant

Presentation for NDIA Systems Engineering
M&S Committee Meeting
August 13, 2009

Insights presented were strongly influenced by work related to the TTD ACTD Modeling Assessment Group (MAG) and by work performed for the JHU Applied Physics Laboratory in support of one of its sponsors.

Note pages identify selected sources and elaborate on some points.

Slide 1

In this presentation, notes pages provide elaboration of points on the slides (such as providing examples) and provide citation information for selected items when such is not on the slide itself.

This presenter was a member of the Modeling Assessment Group (MAG) for the DTRA Tunnel Target Defeat Advanced Concept Technology Demonstration (TTD ACTD). MAG responsibilities and activities (shown below) contributed significantly to Pace's focus on simulation uncertainty issues. The MAG was to:

> "Provide input to the assessment of confidence levels and limitations of the high fidelity codes for targeting applications in terms of:
>
> Accuracy of the test data
>
> Accuracy of modeling capability
>
> Progress in modeling achieved during the TTD ACTD
>
> Model uncertainties
>
> Relative capabilities of the different models
>
> As such, the MAG acts as the VV&A Advisor for the Program Manager"
>
> \- 2005 presentation to TTD ACTD Steering Committee

What Is Uncertainty?

- **National Research Council (NRC) 1994 publication:**

  Uncertainty is "a lack of precise knowledge as to what the truth is, whether qualitative or quantitative" [*Science & Judgment in Risk Assessment*]

- **Pace definition:**

  Simulation-related Uncertainty = anything which prevents or hinders accurate or precise determination of the occurrence, magnitude, duration/continuance, or extent of any factor, parameter, or process associated with a simulation or its products.

Full consideration of uncertainty (however it may be defined) is required for rational use of simulation results.

Slide 2

The fundamental premise of this presentation is that full consideration of uncertainty, however it may be defined, is required for rational use of simulation results. Without full consideration of simulation-related uncertainty, one does not know if conclusions or actions based upon simulation results are sound. This is why the subject is important. While partial consideration of simulation uncertainty is better than ignoring uncertainly totally, partial consideration of uncertainty does not ensure one properly understands simulation results.

This presentation takes a very broad approach to simulation-related uncertainty. It basically is anything that interferes with precise determination of true values.

Distinctions have been made about varieties of uncertainty (aleatory, epistemic, and error, e.g., in AIAA and ASME V&V guides for CFD and computational solid mechanics respectively), but those distinctions are not significant as for as a comprehensive paradigm for simulation-related uncertainty even if they become useful as one tries to reduce or mitigate the impact of uncertainty in a simulation's predictions.

Simulation Uncertainty Needs More Attention
Others Think So Too!

"DoD should seek better methods to characterize, quantify, and manage the uncertainty inherent in all aspects of modeling, simulation, & analysis – including inputs, modeling assumptions, parameters, and options."

Defense Modeling, Simulation, and Analysis: Meeting the Challenge, NRC Committee on Modeling and Simulation for Defense Transformation, 2006

"The national security laboratories and NNSA should expand their use of QMU while continuing to develop, improve, and increase application of the methodology. While they have focused much attention on uncertainty quantification, a broader effort is needed in this area, including further development of the methodology to identify, aggregate, and propagate uncertainties."

Evaluation of Quantification of Margins and Uncertainties Methodology for Assessing and Certifying the Reliability of the Nuclear Stockpile, NRC Committee on the Evaluation of Quantification of Margins and Uncertainties Methodology for Assessing and Certifying the Reliability of the Nuclear Stockpile, 2008.

The fear that explicit (full) treatment of uncertainty might decrease confidence in simulation results was noted by the NRC in 1994. This presenter thinks that attitude still exists.

Chapter 9, "Uncertainty," ***Science and Judgment in Risk Assessment***, Committee on Risk Assessment of Hazardous Air Pollutants, Board on Environmental Studies and Toxicology (BEST), Commission on Life Sciences, National Research Council, The National Academies Press, 1994

Slide 3

Pace is not the only one to recognize that simulation uncertainty needs more attention, as indicated by the quotes from NRC reports.

The two recent NRC reports recommending more attention to simulation uncertainty are significant for the breadth of where they see the need for such attention (DoD) and in a kind of activity that is noted as needing additional attention (QMU). Simulation within the Defense arena tends to be representative of state-of-the-art in M&S. Often more formal attention is given to simulation within DoD than in other domains. So this acknowledgement of the need for additional attention to the issue of uncertainty probably reflects the general situation of all simulation.

QMU has been a central focus of much of the uncertainty quantification (UQ) efforts in the past decade or so. Thus, if QMU needs to give more attention to uncertainty issues, it is likely that there is no area of simulation application which does not also need to give more attention to uncertainty issues.

The problem noted in the mid-1990s of concern that explicit acknowledgment of uncertainty would decrease the impact of simulation results on decision makers still exists as far as this presenter can tell, and it still encourages analysts to fail to treat uncertainty fully and explicitly in their analyses.

Treatment of M&S Uncertainty -- Historical Perspective

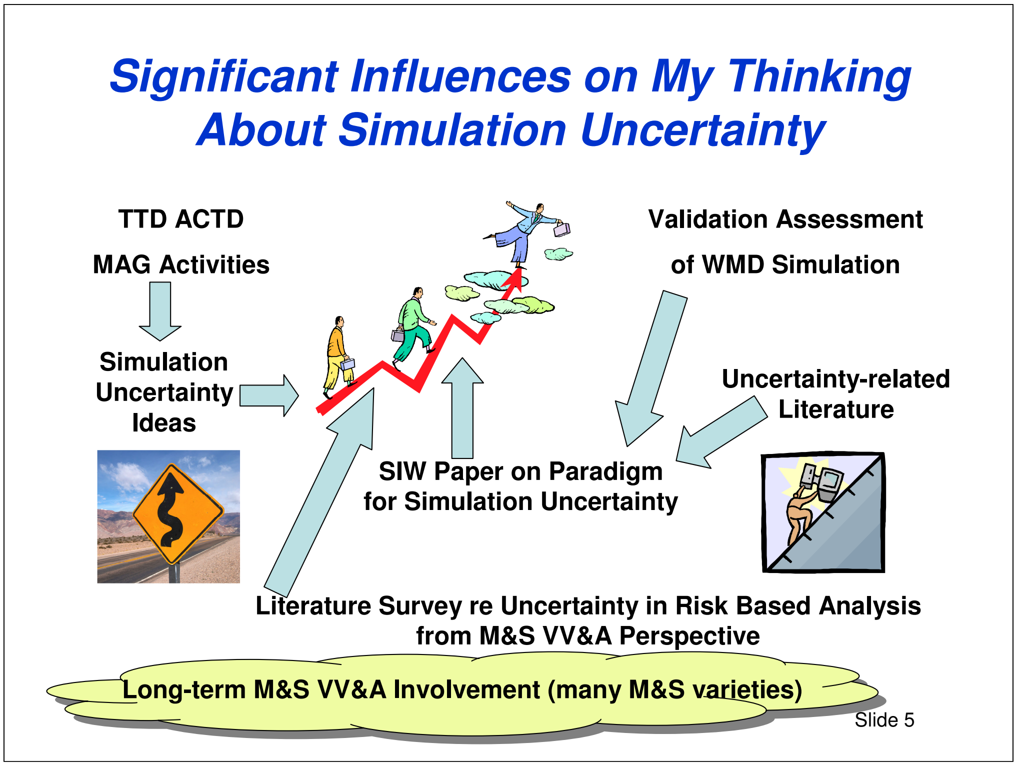
Much Has Been Done, But Not Comprehensively

- **Parameter uncertainty has been the main uncertainty quantification (UQ) focus (Monte Carlo sensitivity studies have been used from M&S earliest days)**
- **Model uncertainty & application uncertainty have been addressed less thoroughly than parameter UQ, although these have been emphasized in past decade or so**
- **Subjective uncertainty in SME estimations & responses to M&S results remains a recognized problem**
- **Many techniques have been identified / developed to address M&S uncertainty (often it is unclear which techniques work best for a situation). Bayesian techniques, evidence theory, etc. have been suggested as ways to expand reliable knowledge subjectively beyond what is available from objective test results.**

Slide 4

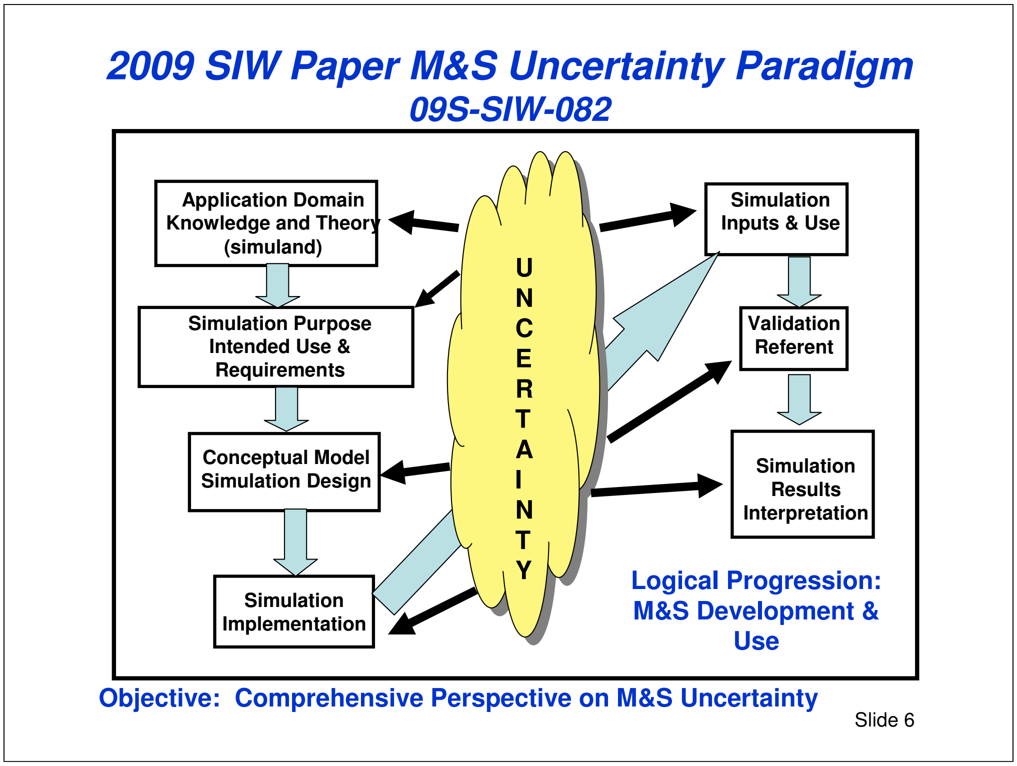
This slide simply indicates a few topics from simulation uncertainty. They are like to tops of ice bergs visible above the water. There is much more unseen (not stated here).

Substantial attention has been given to some aspects of simulation uncertainty. The problem from Pace's perspective is that a total perspective on simulation uncertainty has been lacking. Its absence has made it easier for people to find partial treatment of uncertainty acceptable. The purpose of this presentation is to provide an indication of the scope of total consideration of simulation uncertainty, and to mention the potential magnitude of such uncertainty in some situations so that the importance of the issue will be appreciated.



The SIW paper mentioned on the slide is: Dale K. Pace. "Simulation Uncertainty and Validation," paper 09S-SIW-082 in the Spring 2009 Simulation Interoperability Workshop. Copies available from the Simulation Interoperability Standards Organization (SISO) website at http://www.sisostds.org.

Pace's three decades-plus M&S VV&A involvement has exposed him to uncertainty issues for simulations ranging from a tiny spreadsheet model created with a few months of part-time effort by one staff member to a massive complex simulation with more than 100 staff years invested in simulation development. Simulation implementations have varied from stand-alone simulations in desk top computers to those running on super computers to distributed simulations running on geographically separate computers, and the simulations have included ones with software, hardware, systems, and people in the loop. Applications of the simulations have ranged from first principle physics codes to engineering-level simulations to financial estimators and medical resource tools to computerized support for seminar wargames (which were little more than automated back-of-the-envelope algorithms). Simulation purposes have varied from operational decision support to engineering support to effectiveness analysis to policy analysis. In addition, Pace has been involved in leadership of major M&S VV&A activities (MORS SIMVAL99, Foundations '02, etc.). ***This breadth of background has influenced his perspective on simulation uncertainty and the importance of treating it comprehensively.***



The SIW paper mentioned on the slide is:

Dale K. Pace. “Simulation Uncertainty and Validation,” paper 09S-SIW-082 in the Spring 2009 Simulation Interoperability Workshop. Copies available from the Simulation Interoperability Standards Organization (SISO) website at http://www.sisostds.org.

The logical progression of M&S development and use can be parsed in many ways. Some of the items shown here could be divided into parts and those parts treated separate. Some of the items shown here could be combined. Uncertainty would still impact all of the items, however one chooses to parse the progression.

Whatever paradigm might be used for simulation development, the logical progression shown on this slide occurs. Some development paradigms cycle though steps multiple times during development.

This graphic will be used as a guide to where we are in the presentation as it progresses.

Specific Uncertainty To Be Discussed

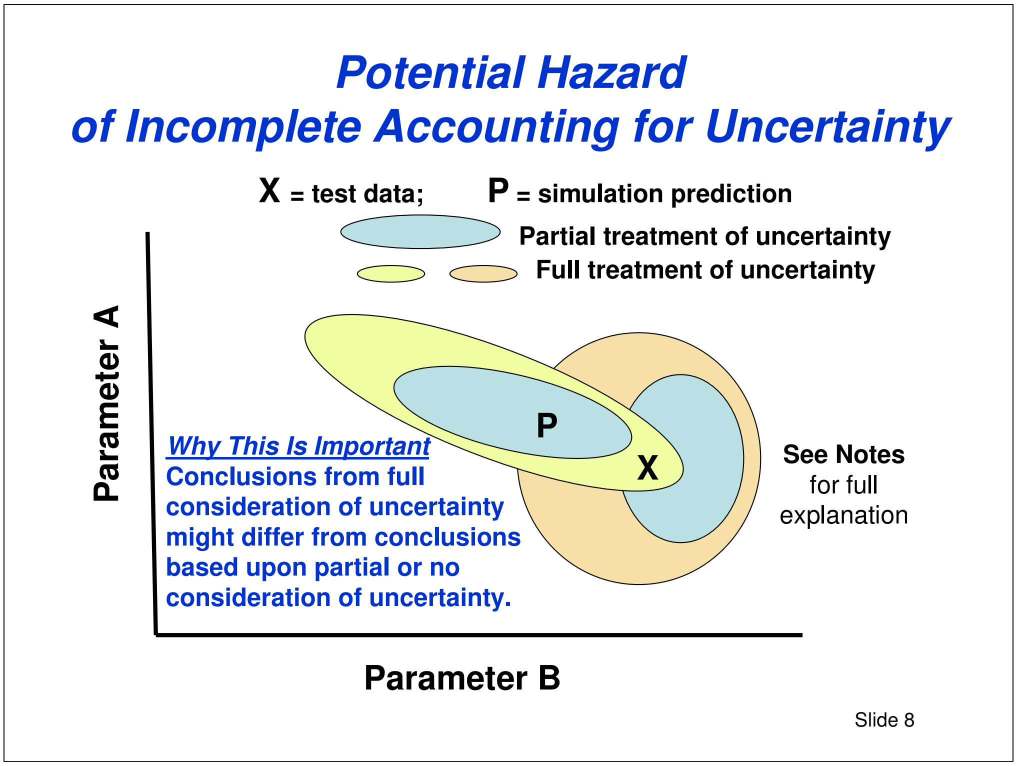
- **Application Domain Uncertainty**
- **Simulation Purpose Uncertainty**
- **Simulation Conceptual Model & Simulation Design Uncertainty**
- **Simulation Implementation Uncertainty**
- **Simulation Input Uncertainty**
- **Simulation User Uncertainty**
- **Simulation Validation Referent Uncertainty**
- **Simulation Results Interpretation Uncertainty**

Input & use uncertainty (together in the paradigm) are treated separately. Each topic is addressed in one or more slides. This is how we will gain a comprehensive perspective on simulation uncertainty.

Slide 7

The way I hope to impart what a comprehensive perspective on simulation uncertainty is by commenting on uncertainty in each of the areas of the paradigm of the previous slide. I separated the simulation input uncertainty from use uncertainty (which is sometimes called "user effects") for presentation convenience.

By addressing uncertainty in each of these areas, I hope to show some of the kinds of things that need to be considered and suggest how significant such considerations may be at times.



In the cartoon, the simulation prediction (P) is quite far from the test result (X). In that case, one probably would conclude that the simulation result was too inaccurate to support its intended use. The same conclusion would probably be reached if there was only a partial accounting for uncertainty, as indicated by the blue areas. These areas indicate where X or P might have occurred (for a specified level of confidence statistically) given the uncertainties considered.

On the other hand, if all uncertainties had been considered (the areas of yellow and tan), then a different conclusion might be reached. Those areas indicate when P and X might have occurred given full consideration of uncertainty. There is considerable overlap of possible simulation predictions and possible test results. In fact, the specific test result (X) is within the simulations potential prediction region; and the simulation prediction (P) is within the region of potential test results.

The point from this chart is simple. Full consideration of uncertainty has the potential to lead to a different conclusion than one might reach by ignoring uncertainty or only considering part of the uncertainty.

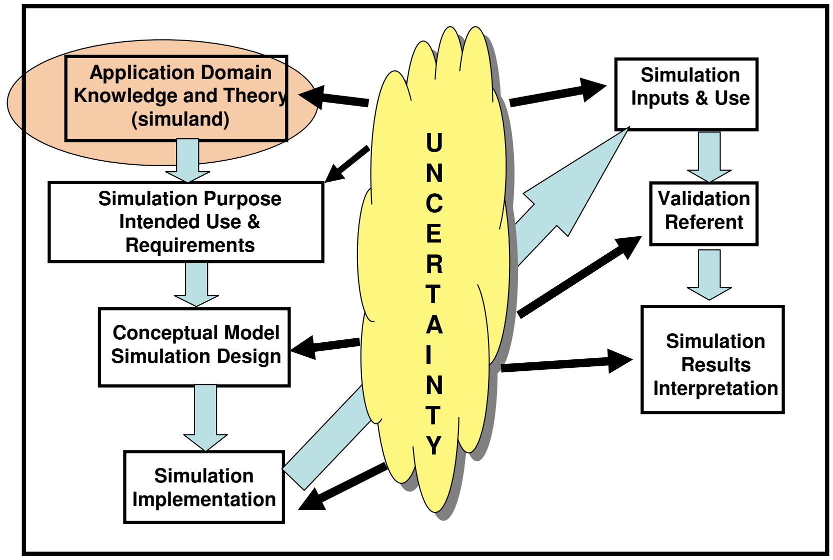
This slide also implies that more sophisticated metrics may be needed to evaluate comparisons of simulation predictions with test results than have been used many times in the past.

Presentation Approach

- **Address each M&S development & use area impacted by uncertainty**
- **Share ideas about uncertainty magnitude, ways to quantify it, and possible ways to control or ameliorate it**
- **Encourage appreciation for importance of treating M&S uncertainty in a comprehensive manner**

Slide 9

Where We Are in Presentation



Slide 10

Application Domain Uncertainty

- **Application domain uncertainty comes from domain knowledge limitations.**
  - **Fundamental knowledge limitations (quantum indeterminacy, Godel's incompleteness theorems).**
  - **Variability of processes, events, objects, etc. in the domain.**
  - **Governing laws and principles not completely understood.**
  - **Domain situation not fully known (and may not be knowable).**
  - **Presumed information about the domain may be incorrect.**
- **Application domain uncertainty varies with both the domain and the application.**
  - **Behavior of individual items in a domain may be unknown even if statistics about such behavior for large number of individual items is well known (e.g., intervals between arrival of items at a location, size of items in a particular class of items).**
  - **Domain information uncertainty may vary with parameter value (e.g., human response to heavy dose of toxic substance may be well known, the person dies, whereas response to a light dose is uncertain).**

Application Domain = M&S Development Referent

Slide 11

There are fundamental limits to what can be known (e.g., quantum indeterminacy and Godel's incompleteness theorems). Even more limiting is our capacity to prove or demonstrate that a computer program will always perform as expected (but that is a simulation implementation uncertainty, not an application domain uncertainty).

Variability of processes, events, objects, etc. of the application domain create uncertainty, even when the variability is well understood and can be described statistically with confidence.

For some application domains, such as ones involving human responses, the laws and principles governing behavior are imperfectly understood and only vaguely captured by computable equations. For other application domains, pertinent laws of physics are well understood, but ability to know the situation fully and correctly is very limited, as in the geology related to ground shock propagation over distances measured in hundreds of meters (joints, their characteristics including orientation and location, are difficult to determine accurately)..

These are the kind of factors that bring uncertainty to the application domain of a simulation.

It helps to think of the application domain as the "M&S development referent." The application domain sets reality bounds on the M&S.

.

Three Varieties of Application Domain Uncertainty

Well Characterized Domains (e.g., production line output):
- **Scientific principles / laws well-known**
- **Scenario / situation fully known**
- **Variability fully characterized and no other uncertainty**
  - **Uncertainty may be small, a few percent or less.**

Semi-Characterized Domains (e.g., ground shock propagation):
- **Scientific principles / laws generally well-known**
- **Scenario / situation partially known**
- **Variability partially characterized and other uncertainty also exists**
  - **Uncertainty may be significant, factor of several.**

Poorly Characterized Domains (e.g., involves human responses):
- **Not all scientific principles / laws are well-known**
- **Scenario / situation is not fully known**
- **Variability poorly characterized and other uncertainty also exists**
  - **Uncertainty may be large, one or more orders of magnitude.**

See notes for elaboration.

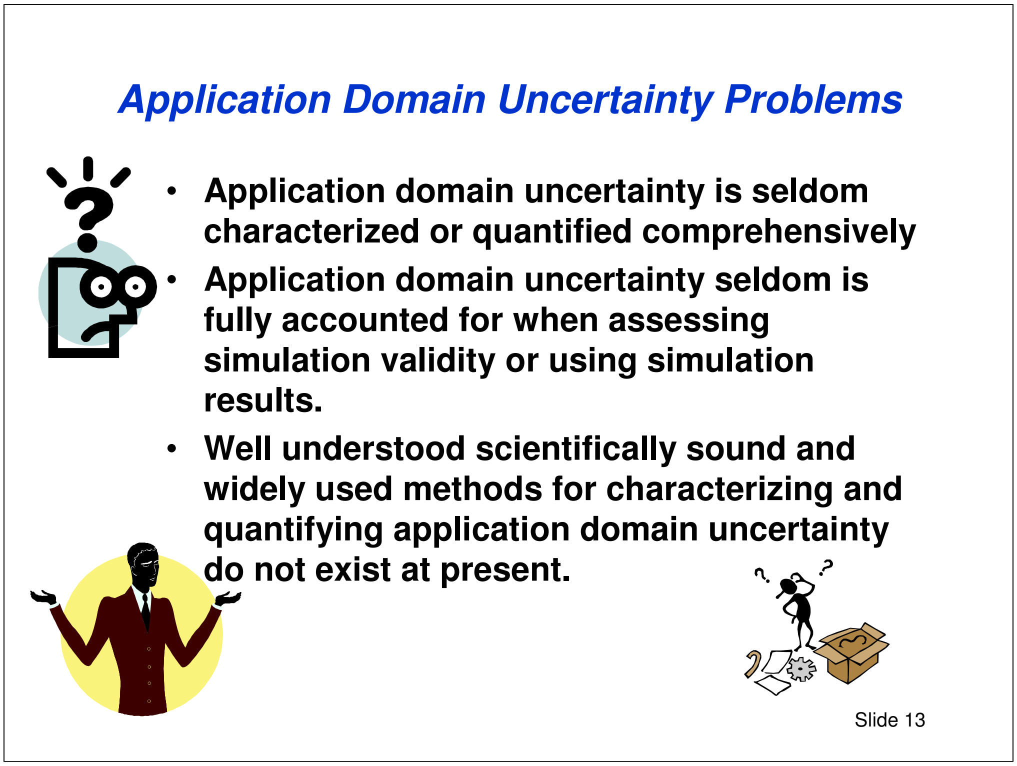
Slide 12

Magnitude and characteristics of application domain uncertainty vary widely. Three general groups are suggested by this slide as a simple way to think about this problem.

The application domain for simulation of a manufacturing production line is an example of a ***well-characterized domain***. Properties of materials used and impact of machining and other processes are well understood so that production line product can satisfy specifications. "Variability" is sometimes called aleatory uncertainty. Reliability engineering has developed methods to address uncertainty in such application domains very competently.

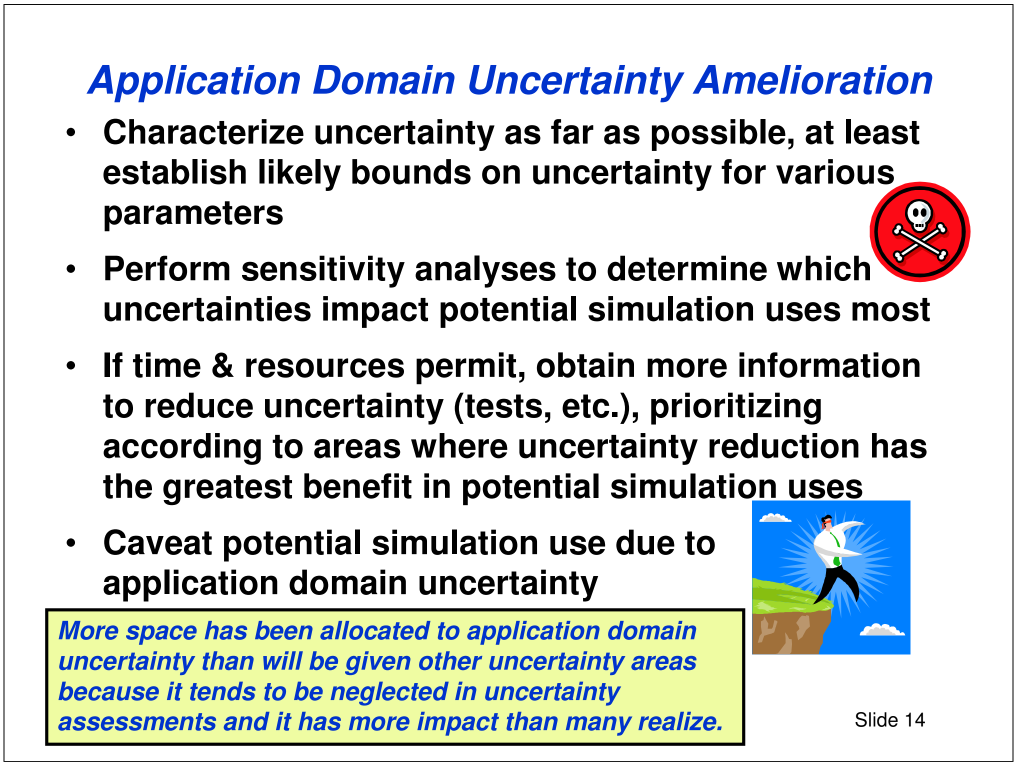
The application domain for simulation of ground shock propagation from an explosion, such as occurred in under ground nuclear tests (UGTs), is an example of a ***semi-characterized domain*** which has substantial uncertainty because of limitations in capability to perceive correctly and fully describe geological conditions. The physics involved are thought to be generally well understood. Scaled parameters (such as peak shock) vs range show a factor of two variation in UGT results [Ch. 4, Effectiveness of Nuclear Weapons Against Hard and Deeply Buried Targets (pp 30-51), NRC 2005 report, *Effects of Nuclear Earth-Penetrator & Other Weapons*]. The variation in results probably is caused mainly by limitations in ability to characterize geology of the sites.

Application domains involving human responses, physically or psychologically, are likely to be ***poorly characterized domains***. E.g.,, a factor of ten is normally used to account for variability in human response to a given dosage of a toxic substance. This is often the spread between the dose level that will affect 5% of a population and the dose level that will affect 95% of that population. Likewise, similar significant differences exist in large-scale test results for transport and diffusion of chemical agent due to meteorological factors and terrain features.



This slide presents three negative conclusions (assertions might be a more appropriate description) that this presenter has reached based upon his awareness of simulation uncertainty literature and simulation use in the U.S. Such negative conclusions (assertions) are difficult to document, and no attempt is made here to justify these points of view by describing the scope of the presenter's knowledge of the literature and simulation community practice.

It is hoped that any who address application domain uncertainty in a substantive way will document that endeavor so that others can benefit from insights obtained by it. Any approach to dealing with simulation uncertainty more comprehensively than is usually done will produce insights for the simulation community if documented. The Sandia report on "Estimation of Total Uncertainty in Modeling & Simulation" identified in a back-up slide illustrates the kind of insight that can come from such documentation.



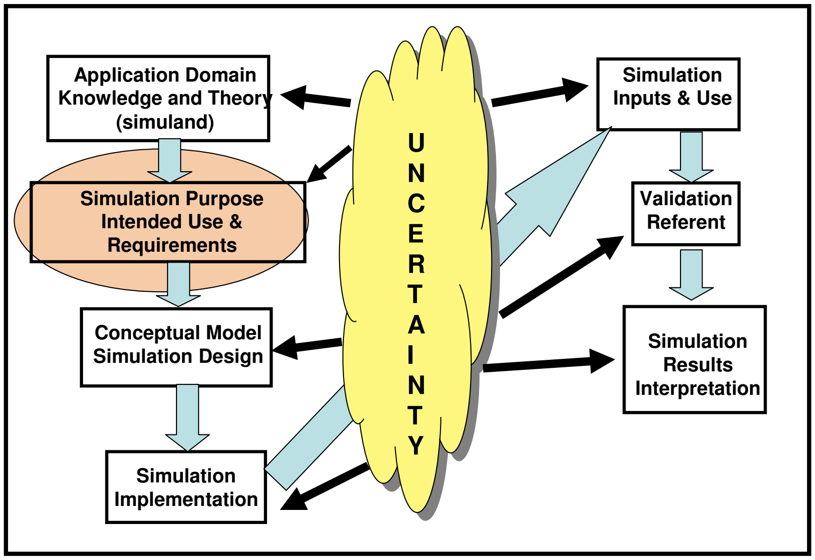
These suggestions, although obvious to a mature simulation practitioner, are seldom done seriously. We need to do these things more often.

Lucas' 2009 thesis (CalTech PhD) on "Uncertainty Quantification Using Concentration-of-Measures Inequalities" describes a method that can help to bound simulation uncertainty.

Sensitivity analysis is a standard technique, but not used as often as needed to identify potential impact of simulation uncertainties.

Caveats about the risk associated with using simulation results is more common now than it had been in the past, but still it is not routine and such caveats do not always address all simulation uncertainty.

Where We Are in Presentation



Slide 15

Simulation Purpose Uncertainty

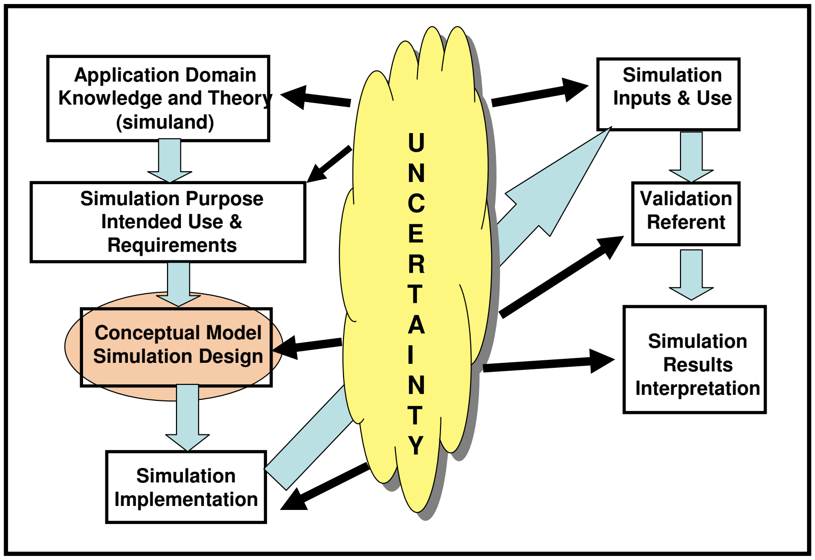
- **Simulation purpose is defined by intended use, requirements, and acceptability criteria**
  - **When done properly, these are explicit, clear (unambiguous), complete, consistent, and correct. Their satisfaction can be demonstrated in assessment.**
  - **It is not uncommon for the items which define simulation purpose to be vague, ambiguous, incomplete, inconsistent, impossible, and even wrong.**
- **Whenever the definition of simulation purpose is not done properly, uncertainty enters the simulation. The uncertainty is hard to quantify, but it can be large.** ***It is hard to hit the target when aiming in the wrong direction.***

Application of requirements engineering insights and thorough V&V of requirements reduces (and may even eliminate) purpose uncertainty potential.

Slide 16

Lack of properly articulated requirements for software developments, which would include simulations, is a major factor in the large number of requirements-related faults (errors) in many software developments. Some attribute the majority of software faults to problems in the requirements (e.g., Lewis, (*Software IV&V*, 1991). Similar comments may be found in more contemporary software engineering literature, such as Ellen Gottesdiener, “Good Practices for Developing User Requirements,” *CrossTalk (J. of Defense Software Engineering), May 2008,* or Steve McConnel, “An Ounce of Prevention,” *IEEE Software*, May / June 2001, that puts correction of requirement errors at about half the software fault repair cost (although the sources he cited were older ones). The same point (the big problem of faulty requirements0 is made by R. N. Charette, “Why Software Fails,” *IEEE Spectrum*, Vol. 42, No. 9 (Sept.2005), pp. 42-49.

Where We Are in Presentation



Slide 17

Simulation Conceptual Model & Simulation Design Uncertainty 1

- **The simulation conceptual model is a mechanism by which simulation requirements are transformed into detailed specifications from which a design can be developed that will fully satisfy the requirements.**
- **Simulation CM/design addresses simulation structure, algorithms, assumptions, processes, etc.**
- **Uncertainty enters via:**
  - **Variability represented by stochastic processes**
  - **Approximations, look-up processes, interpolations, etc.**
  - **Assumptions**
  - **Errors**

Slide 18

Simulation Conceptual Model & Simulation Design Uncertainty 2

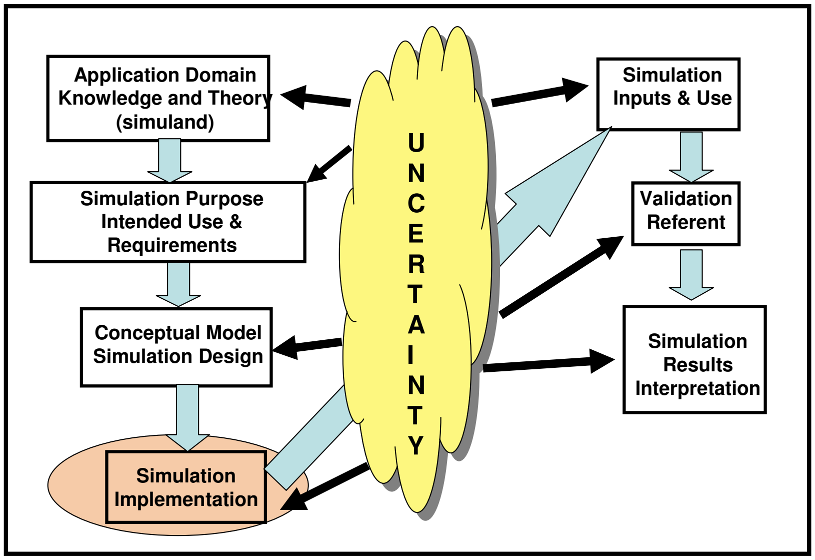
- Risk assessment literature expresses concern about uncertainty (errors) from simulation structure and model errors as well as from other assessment aspects.
- Uncertainty introduced from the CM and simulation design is difficult to quantify, but it has potential to be large. Federate algorithm/assumption incompatibility can be a problem for federations.
- CM & design uncertainties are seldom characterized.
- Potential magnitude of uncertainties here probably decrease with code maturity – errors tend to be discovered and addressed (at least to some degree) with code development and use, although new errors may be introduced with code modifications and "improvements." Corrected errors cause modifications to the conceptual/mathematical model of the code.

Thorough conceptual validation can reduce potential for simulation CM & design uncertainty.

Slide 19

Examples of such model concerns in risk assessment may be found in Bond et al, "Structural Models -- Optimizing Risk Analysis by Understanding Concept Uncertainty," *First Break* (Vol. 26 No. 6, pp. 65-71, 2008), and in Hellstrom, "The science-policy dialogue in transformation: Model-uncertainty and environmental policy," *Science and Public Policy*, Volume 25 (1996), pp. 91-97. Similar concerns about model uncertainty may be found in the text by Fellin et al, *Analyzing Uncertainty in Civil Engineering*, Springer 2005.

Where We Are in Presentation



Slide 20

Simulation Implementation Uncertainty

- **Software bugs and other implementation issues create potential for uncertainty.**
- **Table lookup, grid sizes, and other computational processes also may introduce uncertainty (e.g., different results for the same simulation running the same problem on different operating systems).**
- **In contrast to estimation of undetected software faults, uncertainty in simulation results from software faults is seldom fully characterized or quantified.**

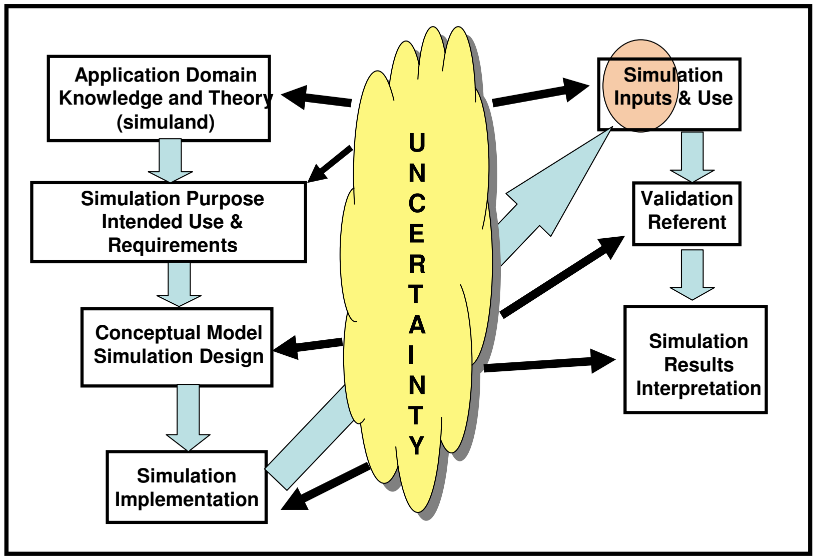
Thorough verification and use of the simulation tend to reduce uncertainty in this area by fault detection and correction.

Slide 21

Nine Causes of Software Errors (from http://www.webdevnotes.com/nine-causes-of-software-error/, accessed 29 June 2009)

- Faulty requirements definition
- Documentation errors
- Client-developer communication failures
- Deliberate deviations from software requirements
- Logical design errors
- Coding errors
- Non-compliance with documentation and coding instructions
- Short-comings of the testing process
- User interface and procedure errors

Where We Are in Presentation



Slide 22

Simulation Input Uncertainty 1

- Garbage In Garbage Out (GIGO) – recognized early in computer simulation history still is a problem.

- Mixing input data from incompatible authoritative sources can create input uncertainties.

- Data entry error can occur, especially for simulations with thousands of inputs, even in the best situations.
- Uncertainties associated with simulation input data are not always fully characterized or quantified.

Careful attention to inputs reduces this kind of uncertainty, but does not eliminate it.

Slide 23

DoD's approach to input uncertainty has been to emphasize authoritative sources for input data, and that has helped a lot but it does not completely solve the problem.

Complex simulations have many inputs. Sometimes multiple authoritative sources are used, especially in a federation whose federates may use different authoritative sources. There may be compatibility/perspective problems in mixing data from multiple sources. Also, parameters available from authoritative sources may not match exactly simulation parameters. This then requires user interpretation or manipulation. This has caused differences of 30-70% when multiple simulations with roughly the same physics representation run exactly the same problem using the same inputs from authoritative sources. It should be noted that this problem could be considered an aspect of "user effects" or of input uncertainty when inputs do not neatly match simulation inputs.

Visualization software and other automation tools can help detect data entry errors before the simulation is run, or in diagnosis of problems after a simulation is run.

Simulation Input Uncertainty 2

- **Input uncertainties can cause code results to be completely wrong. High fidelity simulations are less likely to be used than lower fidelity simulations when inputs are very uncertain.**
- **In tests of high fidelity ground shock propagation codes used to predict damage to underground facilities, Red and Blue descriptions of sites were provided. Comparing differences (10% to a factor of 2) in predications of bottom line parameters from the same code for the Red and Blue site descriptions illustrate the kind of impact that input uncertainty can have on simulation results.**

Input uncertainty causes significant differences in results even with emphasis on input correctness.

Slide 24

In the tests of high fidelity ground shock propagation codes, the Red site descriptions were typical of what might be known about a site to which normal access is denied. Information about it would be based upon general geological information, what might be observed from satellites, information from the Intelligence Community, etc. Blue site descriptions were characteristic of the information that would be available for a site to which one had unrestricted access, such as a site where one expected to perform a test and could observe and measure as needed to obtain information that would be helpful. In both the Red and Blue site descriptions, great care was taken to create correct and as complete as possible site descriptions. The results mentioned indicate that the impact of input uncertainty can be significant for complex simulations even when great care is taken to ensure inputs are correct.

The prediction differences (10% to a factor of 2) were for severe damage predictions for underground facilities, the bottom line parameter. Differences in predictions for intermediate parameters (such as peak particle velocity) could be several times larger.

The only point from the uncertainty quantification mentioned is that the magnitude of uncertainty can be significant. There is no intent to imply that the magnitude of uncertainty mentioned is typical of what one might encounter.

Simulation Input Uncertainty 3

- **In many areas, adequate experimental/test/measured data do not exist and SME estimations are used as inputs or to supplement the data in the inputs.**
  - **Human response to toxic substances (often based upon extrapolation from animal testing, factor of 10 often used to account for interspecies uncertainties)**
  - **River carrying capacity of specific kinds of fish**
  - **Flood risk analysis**
  - **Software development schedule and cost**
  - **Etc.**
- **Techniques exist to reduce (and manage?) SME estimation uncertainty, but in general this uncertainty is poorly characterized and seldom quantified (for individuals or for SME uses).**

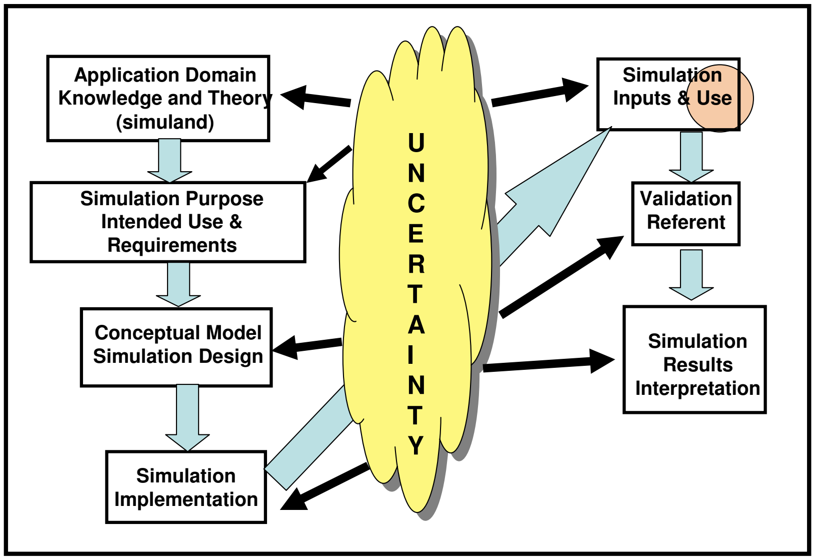
Slide 25

Examples of SME use in the areas noted on the slide are from NRC 1997 report (*Review of Acute Human-Toxicity Estimates for Selected Chemical-Warfare Agents)*, Uusitalo et al ("Estimation of Atlantic Salmon Smolt Carrying Capacity of Rivers Using Expert Knowledge," J. of Marine Science, Vol. 61, No. 4, pp. 708-822, 2005). Merz & Thieken (*Flood Risk Analysis and Uncertainty Considerations*, EGU 2008), and Jorgensen ("A Review of Studies on Expert Estimation of Software Development Effort ," J. of Systems & Software, Vol. 70 Nos. 1-2, pp. 37060, Feb. 2004).

Jorgensen identified and evaluated 12 expert estimation "best practice" guidelines: (1) evaluate estimation accuracy, but avoid high evaluation pressure; (2) avoid conflicting estimation goals; (3) ask the estimators to justify and criticize their estimates; (4) avoid irrelevant and unreliable estimation information; (5) use documented data from previous development tasks; (6) find estimation experts with relevant domain background and good estimation records; (7) Estimate top-down and bottom-up, independently of each other; (8) use estimation checklists; (9) combine estimates from different experts and estimation strategies; (10) assess the uncertainty of the estimate; (11) provide feedback on estimation accuracy and development task relations; and, (12) provide estimation training opportunities.

The DoD M&S VV&A *Recommended Practices Guide* (RPG) contains useful information about SMEs among its Special Topics (http://vva.msco.mil).

Where We Are in Presentation



Slide 26

\

Simulation User Uncertainty 1

- **"User effects" is a term for the impact on code results from choices by the user in problem interpretation and set up (such as the grid chosen for the run)**
- **In simulation literature, user effects impact for hi-fi codes can be reduced from orders of magnitude to factor of two when mitigation is applied (same code applied to same problem by different users), mainly reported by the nuclear engineering community concerned with nuclear power plants (NPPs) and their safety**

When potential user effects are characterized fully and quantified, the magnitude of the uncertainty often is disturbing.

Slide 27

User effects not only show up in significant fashion when the same code is applied to the same problem, but it also shows up when comparable codes tackle the same problem using the same inputs from authoritative sources.

Evidence about the potential magnitude of user effects on code results from multiple users of the same code addressing the same problem range from orders of magnitude differences in results when mitigation efforts are not employed [1], and to at least a factor of two differences after NPP-style mitigation is employed.[2] In the case of multiple users applying the same code to the same problem, the only difference in code results come from user effects. Comparing the differences between code results with different users does not directly indicate how large the impact may be on predictive accuracy since the real result might be between the two predictions or outside them, but it gives an indication of what the magnitude of the impact of user effects might be.

[1] G. Kirchner, S. R. Peterson, S. Bushell, P. Davis, V. Filistovic, T. G. Hinton, P. Krajewski, T. Riesen, and P. U. de Haag, "Effect Of User Interpretation On Uncertainty Estimates : Examples From The Air-To-Milk Transfer Of Radiocesium," Journal of Environmental Radioactivity , 1999 (Vol. 42 No. 2-3), pp. 177-190. Available at http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6VB2-3W5SP39-K&_user=10&_rdoc=1&_fmt=&_orig=search&_sort=d&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=ef20e681d54e1c88b819760987a2db42.

[2] Alessandro Petruzzi and Francesco D'Auria, "Thermal-Hydraulic System Codes in Nuclear Reactor Safety and Qualification Procedures," *Science and Technology of Nuclear Installations*, Volume 2008, Article ID 460795, 16 pages. Available at http://www.hindawi.com/getpdf.aspx?doi=10.1155/2008/460795.

Simulation User Uncertainty 2

- **"User effects" are less of a potential problem when the simulation has fewer options for the user.**
- **Impact of controller decisions in federations which permit controllers to impact federation activities is a special kind of user effect.**
- **Improved user education/training and software modifications to reduce user options are the main ways user effects have been reduced.**

User effects can be significant (20-30% or greater on hi-fi codes) even when world class experts are code users (same problem and inputs, different codes).

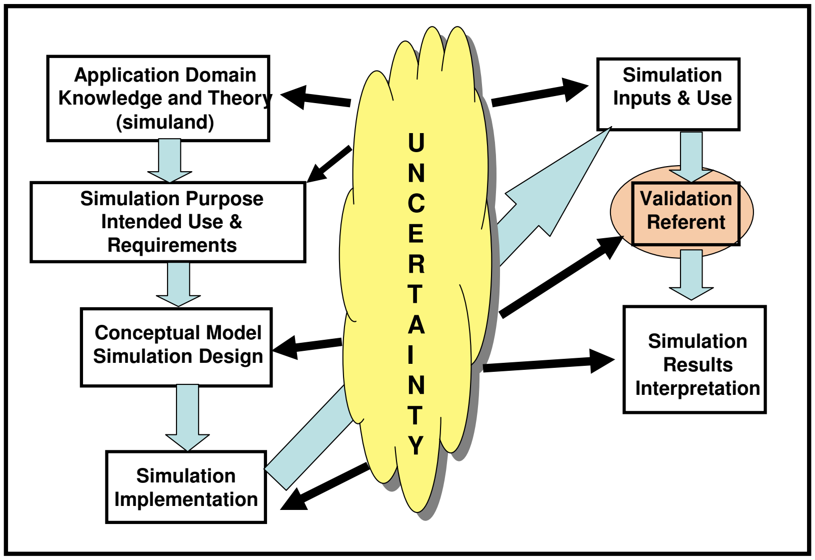
Slide 28

In computer-supported war games, controllers may impact activities of that kind of human in the loop simulation. The uncertainty in simulation results can be significant.

A classic example of uncertainty introduced by simulation controllers comes from the way a Japanese admiral over-ruled umpire assessments in war game evaluation of possible outcomes from the Battle of Midway in WWII. His actions caused basically correct assessments to become totally wrong.

Different codes with roughly the same physics representation were applied to identical problems (same information for inputs). That the several codes considered contained roughly the same physics had been demonstrated in a series of verification exercises and small scale tests. Differences in code prediction (which were a results of user effects and any differences in the codes) were non-trivial, even though the code users were extremely knowledgeable of the codes and true experts in the application domain.

Where We Are in Presentation



Slide 29

Simulation Validation Referent Uncertainty

- **Validation referent inherits the uncertainty of the application domain. It may compound that uncertainty by measurement uncertainty for tests from which data are drawn, and by uncertainty associated with SMEs who may be used in subjective aspects of the referent.**
- **Referent data for comparison with code predictions come in various forms:**
  - **Test results not known before predictions**
  - **Test results known before predictions**
  - **Results from other codes**
  - **Engineering (SME) judgment & theory**
  - **Combination of two or more of these**

Referent data uncertainties should be identified, characterized, and quantified - which often is not done.

Slide 30

In estimating code predictive accuracy, code results have to be compared with referent data. There are a variety of potential referent data, if the term "data" is used loosely. Engineering judgment is sometimes used as referent data when judging code outputs. Correspondence with accepted theory is sometimes used. However, confidence is low when estimations of code predicted accuracy are based only upon comparison of code results with SME engineering judgment, accepted theory, or results of other simulations.

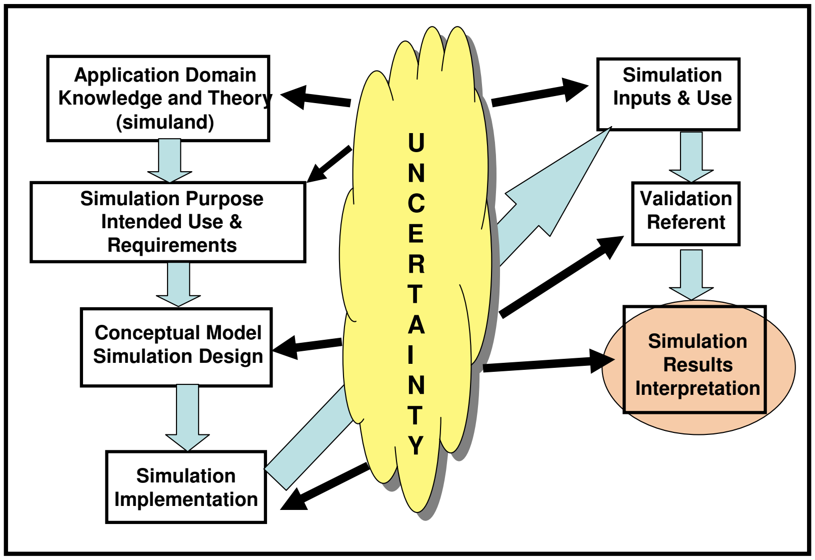
Sometimes code results are compared with test data that the user knows before the code is run. Comparison of code predictions with test results known to the code user prior to running the problem (sometimes called "a posteriori comparisons") is often used in what is called model/code "calibration." Calibration is the process of revising code parameters to bring code predictions into better agreement with test results. This revision process of calibration is also "commonly known as *model updating*, *model tuning*, *parameter calibration*, and *parameter estimation*."[ANSI Standard V&V 10, ASME Guide for Verification and Validation in Computational Solid Mechanics, 29 December 2006] However, many would not consider such calibration efforts as demonstration of a code's validation or predictive accuracy.

The preferred referent data for determining a code's predictive accuracy is test data which the user does not know prior to running the code. Unfortunately, such data may not be available in adequate quantity or quality (conditions related to the test may not be known well enough for clear understanding of uncertainties associated with the data) to permit strong conclusions about code predictive accuracy.

The Uncertainty Quantification (UQ) literature and AIAA/ASME V&V guides emphasis including uncertainties of the experimental data used in validation assessment of codes; they focus on the immediate uncertainties of the particular tests that such data may be drawn from and seldom address the larger issues of uncertainties in the application domain.

This is better than ignoring uncertainties of the validation referent, but less than the comprehensive perspective needed.

Where We Are in Presentation



Slide 31

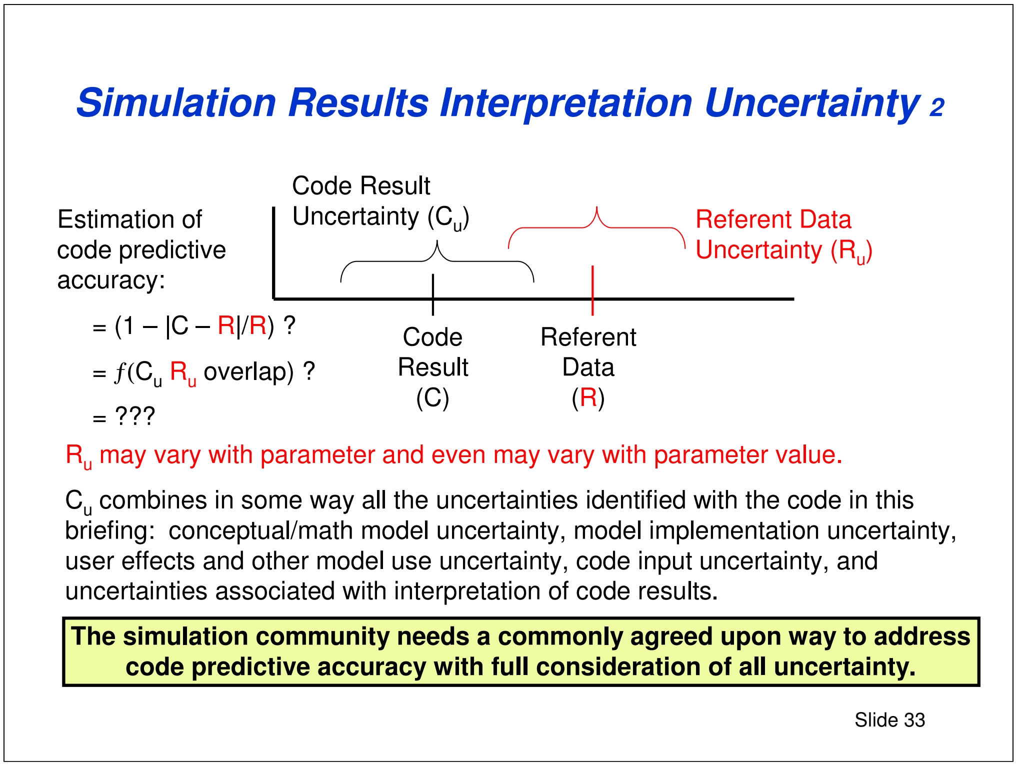
Simulation Results Interpretation Uncertainty 1

- **SMEs are often involved in simulation results interpretation, which introduces many of the uncertainties associated with SME use:**
  - **SME perspective re simulation intended use**
  - **SME bias or slant on the subject**
  - **Differences among SMEs if more than one is involved**
- **Data sampling (M&S predictions and referent) used can impact conclusions significantly.**
- **When uncertainty in both referent and simulation results is considered, more sophisticated processes are required for assessment and conclusion than has been done when such uncertainty is not addressed fully**
- **Risk attitude (risk adverse, risk neutral, risk taker) strongly influences how simulation results may be interpreted**

Slide 32

The personal/subjective aspect of uncertainty for results interpretation is a major concern, and seldom is addressed in a quantitative manner.

The sampling issue is also a source of uncertainty. Careless selection of what part of the application domain is tested (either with referent data or by simulation predictions) can cause significant uncertainty about M&S capability elsewhere in the application domain.



The situation is more complex than shown on this slide. It is multi-dimensional as was illustrated by the two dimension cartoon shown early in the presentation.

Referent data have uncertainties, whether test measurements, results from other codes, theory, or expect opinion. How to identify, characterize, and quantify such uncertainties can be challenging. For test results, data have to be reviewed carefully, removing bad data (but only bad data, no good data should be removed). Then it is necessary to try to determine errors and uncertainties associated with the data that are considered good (valid).

Code uncertainty likewise has to be identified, characterized, and quantified. This briefing has presented various bits of information that suggest something about the magnitude of different simulation aspects. Conclusions about the most appropriate ways to combine the various aspects of code uncertainties have not yet been reached.

Once referent data and code uncertainties have been addressed, the question of how to estimate code predictive accuracy must be faced. The simple approach most often employed to date has been to compute the percent error in the code predictions and subtract that from one to get code predictive accuracy, which usually is presented without reference to uncertainty. That approach probably no longer satisfies the risk assessment perspective frequently found in contemporary materials for code V&V. All uncertainties have to be considered and addressed.

Conclusions

- **The spectrum of aspects that need consideration for full assessment of simulation uncertainty has been indicated. It has been noted that this is seldom done.**
- **Characterization and quantification of uncertainty is well established in some areas, but not for other areas.**
- **Processes to reduce simulation uncertainty involve disciplines like requirements engineering, V&V, etc.**
- **Standard statistical processes, including those incorporating subjective judgment, can help characterize simulation uncertainties.**
- **How best to communicate uncertainty to users is not determined**
- **When simulation uncertainty is fully assessed, its magnitude probably will be disturbing.** ***This poses a challenge to simulationists and analysts: to speak with candor about simulation results and their uncertainty.***

Slide 34

This slide has my bottom line. I think they are valid conclusions from what I presented.

Some may be disappointed that I did not present a table of the potential uncertainty range for the different areas where uncertainty can impact simulation. The software engineering literature produces that kind of information based upon data from hundreds of case histories. We do not have such data for enough simulations to make that kind of table possible. It would be great if in the future that kind of information were to exist. Perhaps if each of us strives to document simulation uncertainty within our spheres of influence, that kind of information about simulation uncertainty will become available.

The candor challenge confronts every knowledgeable user of simulation. We know that "decision makers" often do not want fuzzy guidance from simulation results. They want answers. Sometimes evaluation of us as professionals will be negatively impacted by candor. I hope we always have the courage and integrity to tell it like it really is. That includes a full accounting for simulation uncertainty.

Backup slides have a few items of possible interest.

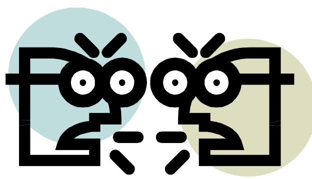
Pace's Concluding Comments

- **My objective: stimulate interest in comprehensive treatment of uncertainty in simulation**
- **What I did:**
  - **Provided a strawman paradigm for comprehensive treatment of simulation uncertainty**
  - **Suggested uncertainty is often larger than we like**
  - **Indicated that we need much better methods for dealing with some aspects of uncertainty**
- **What I hope will occur:**
  - **Members of the NDIA M&S Committee will begin to think about and publish ideas on comprehensive consideration of simulation uncertainty**
  - **All will insist on comprehensive treatment of uncertainty in simulations with which we are involved**

Slide 35

My emphasis upon V&V as mechanisms to ameliorate uncertainty should be expected, given my background and interests, but I think the points made about V&V relative to uncertainty are valid in spite of any bias I may have in that regard.

Your Questions & Comments



Backup Slides

Slide 37

Aspects of Uncertainty

- **Code V&V guides for computational science and engineering suggests three aspects to code uncertainty**
  - ***Aleatory uncertainty*** **(sometimes called *variability*) represents the inherent variation associated with the physical system or environment under consideration, and usually is stochastic.**
  - ***Epistemic uncertainty*** **is potential deficiency in the accurate correct representation of the phenomena due to lack of knowledge. Epistemic uncertainties often have subjective elements.**
  - ***Error.*** **This approach to uncertainty permits *error* to be distinguished from aleatory and epistemic uncertainties by defining error as a recognizable deficiency.**

Slide 38

The threefold perspective on uncertainty presented here follows an approach that is relatively widely used in the computational science and engineering simulation community, and as such seems pertinent to TTD applications by high fidelity codes. The approach has influenced simulation verification and validation (V&V) guides[1] from AIAA and ASME as well as V&V approaches employed within the Department of Energy's Advance Simulation & Computing (ASC) program (previously known as the Accelerated Strategic Computing Initiative, ASCI) and elsewhere.[2] Oberkampf et al (2000)[3] provide background and rationale for this perspective. In this approach:

[1] ANSI Standard V&V 10, ASME Guide for Verification and Validation in Computational Solid Mechanics, 29 December 2006, and Guide for the Verification and Validation of Computational Fluid Dynamics Simulations, AIAA G-077-1998, American Institute of Aeronautics and Astronautics, Reston, Va.

[2] For examples, Timothy G. Trucano, Martin Pilch, and William L. Oberkampf, *General Concepts for Experimental Validation of ASCI Code Applications*, Sandia Report SAND2002-0341, March 2002. Available at http://csdl.ics.hawaii.edu/~johnson/hpcs/doc/Trucano02.pdf, and the NASA Technical Standard NASA-STD-7009 (11 July 2008) *Standard for Models and Simulations* cites the terminology of the AIAA and ASME V&V guides for its uncertainty quantification aspects.

[3] William L. Oberkampf, Sharon M. DeLand, Brian M. Rutherford, Kathleen V. Diegert, and Kenneth F. Alvin, *Estimation of Total Uncertainty in Modeling and Simulation,* Sandia Report SAND2000-0824, April 2000. Available at http://www.prod.sandia.gov/cgi-bin/techlib/access-control.pl/2000/000824.pdf.

A Sampling of Endeavors with a Broad View of Simulation Uncertainty [1]

Estimation of Total Uncertainty in Modeling & Simulation (SAND2000-0824 by Oberkampf, DeLand, Rutherford, Diegert, and Alvin).

> "A comprehensive, new view of the general phases of modeling and simulation is proposed, consisting of the following phases: conceptual modeling of the physical system, mathematical modeling of the conceptual model, discretization and algorithm selection for the mathematical model, computer programming of the discrete model, numerical solution of the computer program model, and representation of the numerical solution. . . . In each of these phases, general sources of variability, uncertainty, and error are identified."

> ***This Sandia report is an excellent piece of work except it does not quite treat all aspects of simulation uncertainty. For example, some aspects of "user effects" seem not to have been addressed.***

Slide 39

This 2000 Sandia report by William L. Oberkampf, Sharon M. DeLand, Brian M. Rutherford, Kathleen V. Diegert, and Kenneth F. Alvin addresses the topics it deals with in a thorough and instructive manner. The level of detail with which the subject is discussed is very helpful.

While the issue of variability in parameters and factors is addressed as well as error, it is unclear that the methodology presented fully addresses uncertainties in the application domain knowledge or fully accounts for all aspects of user effects.

A Sampling of Endeavors with a Broad View of Simulation Uncertainty $_2$

Academic simulation uncertainty perspective extensions

PhD Theses:

- Arthur Peterson, *Simulating Nature: A Philosophical Study of Computer-Simulation Uncertainties and Their Role in Climate Science and Policy Advice*, PhD Dissertation, Vrije Universiteit, Holland, 2006. **His paradigm stimulated the one that presented in this briefing.**
- Krishna Kanta Panthi, *Analysis of Engineering Geological Uncertainties Related to Tunnelling in Himalayan Rock Mass Conditions, Doctoral Thesis, Norwegian Unicersity of Science and Technology, 2006.*

Historical Review:

- Herbert H. Einstein, "Uncertainty in Rock Mechanics and Rock Engineering – Then and Now," ISRM 2003 – Technology Roadmap for Rock Mechanics, South African Institute of Mining and Metallurgy, 2003.

Slide 40

Each of the three items on this chart takes a broader view of simulation uncertainty than is often done.

Peterson approached the subject of simulation uncertainty philosophically, and thereby had a very broad perspective. The structure by which he organized his discussion of simulation uncertainty is very similar to the structure of the paradigm of this presentation.

A Sampling of Endeavors with a Broad View of Simulation Uncertainty $_3$

A Comprehensive Paradigm for Simulation Uncertainty

"Simulation Validation and Uncertainty," Paper 082 at Simulation Interoperability Standards Organization (SISO) Spring Simulation Interoperability Workshop (SIW), March 2009.

- Paradigm developed to support validation assessment of model supporting deliberate planning for military operations in WMD
- The paradigm was used to provide structure for uncertainty analysis addressed in this presentation